FORM PTO-1390 U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371		ATTORNEY'S DOCKET NUMBER **PB-0005 USN** U.S. APPLICATION NO. (If known, see 37 CFR 1.5) **TO BE ASSIGNED** 09/830913
INTERNATIONAL APPLICATION NO. **PCT/US99/25457**	INTERNATIONAL FILING DATE **28 October 1999**	PRIORITY DATE CLAIMED **06 November 1998**

TITLE OF INVENTION
CORTICOSTEROID SYNTHESIS-ASSOCIATED GENES

APPLICANT(S) FOR DO/EO/US
INCYTE PHARMACEUTICALS, INC.; WALKER, Michael G.; VOLKMUTH, Wayne; KLINGLER, Tod M.

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is the **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☐ This is an express request to promptly begin national examination procedures (35 U.S.C. 371 (f)).
4. ☐ The US has been elected by the expiration of 19 months from the priority date (PCT Article 31).
5. ☒ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   a. ☐ is attached hereto (required only if not communicated by the International Bureau)
   b. ☐ has been communicated by the International Bureau.
   c. ☒ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).
7. ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
   a. ☐ are attached hereto (required only if not communicated by the International Bureau).
   b. ☐ have been communicated by the International Bureau.
   c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   d. ☒ have not been made and will not be made.
8. ☐ An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. ☐ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. ☐ An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11 to 16 below concern document(s) or information included:

11. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.27 and 3.31 is included.
13. ☐ A FIRST preliminary amendment.
    ☐ A SECOND or SUBSEQUENT preliminary amendment.
14. ☐ A substitute specification.
15. ☐ A change of power of attorney and/or address letter.
16. ☒ Other items or information:

1) Transmittal Letter (2 pp, in duplicate)
2) Return Postcard
3) Express Mail Label No.: **EL 856 113 067 US**
4) Request to Transfer

09830913 .051302

JC18 Rec'd PCT/PTO 0 2 MAY 2001

U.S. APPLICATION NO. (if known, see 37 CFR 1.5) TO BE ASSIGNED 09/830913	INTERNATIONAL APPLICATION NO.: PCT/US99/ 25457	ATTORNEY'S DOCKET NUMBER PB-0005 USN

17. ☒ The following fees are submitted:

BASIC NATIONAL FEE (37 CFR 1.492(a)(1)-(5):

Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO.....$1000.00

☐International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO..$860.00

International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO.....$710.00

☒**International preliminary examination fee paid to USPTO (37 CFR 1.482) but all claims did not satisfy provisions of PCT Article 33(1)-(4).....$690.00**

☐International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4).....$100.00

ENTER APPROPRIATE BASIC FEE AMOUNT =				$690.00	
Surcharge of **$130.00** for furnishing the oath or declaration later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492(e)).				$	
CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE		
Total Claims	14 =	0	X **$ 18.00**	$	
Independent Claims	2 =	0	X **$ 80.00**	$	
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+ **$270.00**	$	
TOTAL OF ABOVE CALCULATIONS =				$690.00	
☐ Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2.				$	
SUBTOTAL =				$690.00	
Processing fee of **$130.00** for furnishing the English translation later than ☐ 20 ☐ 30 months from the earliest clailmed priority date (37 CFR 1492(f)). +				$	
TOTAL NATIONAL FEE =				$690.00	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by the appropriate cover sheet (37 CFR 3.28, 3.31). **$40.00** per property +				$	
TOTAL FEES ENCLOSED =				$690.00	
				Amount to be Refunded:	$
				Charged:	$

a. ☐ A check in the amount of $_____ to cover the above fees is enclosed.

b. ☒ Please charge my Deposit Account No. 09-0108 in the amount of $ 690.00 to cover the above fees.

c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 09-0108. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

INCYTE GENOMICS, INC.
3160 Porter Drive
Palo Alto, CA 94304

SIGNATURE

NAME: Diana Hamlet-Cox

REGISTRATION NUMBER: 33,302

DATE: 1 May 2001

CORTICOSTEROID SYNTHESIS-ASSOCIATED GENES

TECHNICAL FIELD

The invention relates to seven corticosteroid synthesis-associated genes identified by their coexpression with known corticosteroid synthesis genes; to their corresponding polypeptides; to the use of these biomolecules in diagnosis, prognosis, prevention and evaluation of therapies for diseases, particularly for diseases associated with corticosteroid synthesis or steroid imbalance.

BACKGROUND ART

Steroid hormones such as progesterone, pregnenolone, corticosterone, aldosterone, testosterone, and estrogen play critical roles in reproductive medicine, cardiovascular disease, breast cancer, prostate cancer, osteoporosis, diabetes, and menopausal symptoms (Pavlik (1997) Estrogens, progestins, and their antagonists, Berkhauser, Boston MA, pp. 3-176; Goldfein, In: Katzung, (1995) Basic and clinical pharmacology, Appleton & Lange, Norwalk CT, pp. 592-607; Laycock and Wise (1996) Essential Endocrinology, Oxford University Press, London UK; and Norman and Litwack (1997) Hormones, Academic Press, San Diego CA). Many genes that participate in and regulate steroid synthesis are known, but many remain to be identified. Identification of additional genes will provide new diagnostic and therapeutic targets.

The present invention provides new compositions that are useful for diagnosis, prognosis, treatment, prevention, and evaluation of therapies for cardiovascular disease, breast cancer, prostate cancer, osteoporosis, diabetes, and menopausal symptoms, and for reproductive medicine applications such as contraception and infertility.

DISCLOSURE OF THE INVENTION

In one aspect, the invention provides for a substantially purified polynucleotide comprising a gene that is coexpressed with one or more known corticosteroid synthesis genes in a plurality of biological samples. Preferably, known corticosteroid synthesis genes are selected from the group consisting of steroid acute regulatory gene, P450scc cholesterol side-chain cleavage enzyme, 3-beta-hydroxysteroid dehydrogenase, Type I 3-beta-hydroxysteroid dehydrogenase, Type II 3-beta-hydroxysteroid dehydrogenase, P450c11 beta-hydroxylase, and P450c17 alpha-hydroxylase. Preferred embodiments include (a) a polynucleotide sequence selected from the group consisting of SEQ ID NOs:1-7; (b) a polynucleotide sequence which encodes the polypeptide sequence of SEQ ID NO:8 or 9; (c) a polynucleotide sequence having at least 70% identity to the polynucleotide sequence of (a) or (b); (d) a polynucleotide sequence comprising at least 10, preferably at least 18, sequential nucleotides of the polynucleotide sequence of (a), (b), or (c); (e) a polynucleotide sequence which is complementary to the polynucleotide sequence of (a), (b),(c), or (d); and (f) a polynucleotide which hybridizes under stringent conditions to the polynucleotide of (a), (b), (c), (d) or (e).

Furthermore, the invention provides an expression vector comprising any of the above described polynucleotides and host cells comprising the expression vector. Still further, the invention provides a method for treating or preventing a disease or condition associated with the altered expression of a gene that is coexpressed with one or more known corticosteroid synthesis genes comprising administering to a subject in need a polynucleotide described above in an amount effective for treating or preventing said disease.

In a second aspect, the invention provides a substantially purified polypeptide comprising the gene product of a gene that is coexpressed with one or more known corticosteroid synthesis genes in a plurality of biological samples. The known corticosteroid synthesis gene may be selected from the group consisting of steroid acute regulatory gene, P450scc cholesterol side-chain cleavage enzyme, 3-beta-hydroxysteroid dehydrogenase, Type I 3-beta-hydroxysteroid dehydrogenase, Type II 3-beta-hydroxysteroid dehydrogenase, P450c11 beta-hydroxylase, and P450c17 alpha-hydroxylase. Preferred embodiments are (a) a polypeptide sequence of SEQ ID NO: 8 or 9; (b) a polypeptide sequence having at least 85% identity to the polypeptide sequence of (a); and (c) a polypeptide sequence comprising at least 6 sequential amino acids of the polypeptide sequence of (a) or (b).

In another aspect, the invention provides a pharmaceutical composition comprising a polynucleotide of or a polypeptide in conjunction with a suitable pharmaceutical carrier and a method for treating or preventing a disease or condition associated with the altered expression of a gene that is coexpressed with one or more known corticosteroid synthesis genes comprising administering to a subject in need such a composition in an amount effective for treating or preventing said disease.

In a further aspect, the invention provides a ribozyme that cleaves a polynucleotide of the invention and a method for treating or preventing a disease or condition associated with the increased expression of a gene that is coexpressed with one or more known corticosteroid synthesis genes. The method comprises administering to a subject in need the ribozyme in an amount effective for treating or preventing said disease.

In yet a further aspect, the invention provides a method for diagnosing a disease or condition associated with the altered expression of a gene that is coexpressed with one or more known corticosteroid synthesis genes wherein each known corticosteroid synthesis gene is selected from the group consisting of steroid acute regulatory gene, P450scc cholesterol side-chain cleavage enzyme, 3-beta-hydroxysteroid dehydrogenase, Type I 3-beta-hydroxysteroid dehydrogenase, Type II 3-beta-hydroxysteroid dehydrogenase, P450c11 beta-hydroxylase, and P450c17 alpha-hydroxylase. The method comprises the steps of (a) providing a sample comprising one or more of said coexpressed genes; (b) hybridizing a polynucleotide to said coexpressed genes under conditions effective to form one or more hybridization complexes; (c) detecting the hybridization complexes; and (d) comparing the levels of the hybridization complexes with the level of hybridization complexes in a non-diseased sample, wherein

altered expression levels indicate the presence of the disease or condition.

Additionally, the invention provides antibodies that bind specifically to any of the above described polypeptides and a method for treating or preventing a disease or condition associated with the altered expression of a gene that is coexpressed with one or more known corticosteroid synthesis genes comprising administering to a subject in need such an antibody in an amount effective for treating or preventing said disease.

BRIEF DESCRIPTION OF THE SEQUENCE LISTING

The Sequence Listing provides exemplary corticosteroid synthesis-associated sequences including polynucleotide sequences, SEQ ID NOs:1-7, and polypeptide sequences, SEQ ID NOs:8 and 9. Each sequence is identified by a sequence identification number (SEQ ID NO) and by the Incyte Clone number from which the sequence was first identified.

MODES FOR CARRYING OUT THE INVENTION

It must be noted that as used herein and in the appended claims, the singular forms "a", "an", and "the" include the plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

DEFINITIONS

"NSEQ" refers generally to a polynucleotide sequence of the present invention, including SEQ ID NOs:1-7. "PSEQ" refers generally to a polypeptide sequence of the present invention, including SEQ ID NOs:8 and 9.

A " variant" refers to either a polynucleotide or a polypeptide whose sequence diverges from SEQ ID NOs:1-7 or SEQ ID NOs:8 and 9, respectively. Polynucleotide sequence divergence may result from mutational changes such as deletions, additions, and substitutions of one or more nucleotides; it may also occur because of differences in codon usage. Each of these types of changes may occur alone, or in combination, one or more times in a given sequence. Polypeptide variants include sequences that possess at least one structural or functional characteristic of SEQ ID NOs:8 and 9.

A "fragment" can refer to a nucleic acid sequence that is preferably at least 20 nucleic acids in length, more preferably 40 nucleic acids, and most preferably 60 nucleic acids in length, and encompasses, for example, fragments consisting of nucleic acids 1-50 of SEQ ID NOs:1-7. A "fragment" can also refer to polypeptide sequences which are preferably at least 5 to about 15 amino acids in length, most preferably at least 10 amino acids long, and which retain some biological or immunological activity of a protein sequence, such as SEQ ID NO:8 or 9.

"Gene" or "gene sequence" refers to the partial or complete coding sequence of a gene. The term also refers to 5' or 3' untranslated regions of a transcript. The gene may be in a sense or antisense (complementary) orientation.

"Known corticosteroid synthesis gene" refers to a gene sequence which has been previously identified as useful in the diagnosis, treatment, prognosis, or prevention of diseases associated with corticosteroid synthesis. Typically, this means that the known gene is expressed at higher levels in tissue abundant in known corticosteroid synthesis transcripts when compared with other tissue.

"Corticosteroid synthesis-associated gene" refers to a gene sequence whose expression pattern is similar to that of the known corticosteroid synthesis genes and which are useful in the diagnosis, treatment, prognosis, or prevention of diseases associated with corticosteroid synthesis.

"Substantially purified" refers to a nucleic acid or an amino acid sequence that is removed from its natural environment and is isolated or separated, and is at least about 60% free, preferably about 75% free, and most preferably about 90% free from other components with which it is naturally present.

"Complementary" refers to a sequence having sufficient sequence identity to another sequence so that it will form a stable duplex with that other sequence or a triplex including that other sequence.

THE INVENTION

The present invention encompasses a method for identifying biomolecules that are associated with a specific disease, regulatory pathway, subcellular compartment, cell type, tissue type, or species. In particular, the method identifies gene sequences useful in diagnosis, prognosis, treatment, prevention, and evaluation of therapies for diseases associated with corticosteroid synthesis, particularly diseases associated with corticosteroid synthesis or steroid imbalance; and to the use of these biomolecules in other aspects of reproductive medicine including contraception and fertility.

The method entails first identifying polynucleotides that are expressed in a plurality cDNA libraries. The identified polynucleotides include genes of known function, genes known to be specifically expressed in a specific disease process, subcellular compartment, cell type, tissue type, or species. Additionally, the polynucleotides include genes of unknown function. The expression patterns of the known genes are then compared with those of the genes of unknown function to determine whether a specified coexpression probability threshold is met. Through this comparison, a subset of the polynucleotides having a high coexpression probability with the known genes can be identified. The high coexpression probability correlates with a particular coexpression probability threshold which is less than 0.001, and more preferably less than 0.00001.

The polynucleotides originate from cDNA libraries derived from a variety of sources including, but not limited to, eukaryotes such as human, mouse, rat, dog, monkey, plant, and yeast and prokaryotes such as bacteria and viruses. These polynucleotides can also be selected from a variety of sequence types

including, but not limited to, expressed sequence tags (ESTs), assembled polynucleotide sequences, full length gene coding regions, introns, regulatory sequences, 5' untranslated regions, and 3' untranslated regions. To have statistically significant analytical results, the polynucleotides need to be expressed in at least three cDNA libraries.

The cDNA libraries used in the coexpression analysis of the present invention can be obtained from blood vessels, heart, blood cells, cultured cells, connective tissue, epithelium, islets of Langerhans, neurons, phagocytes, biliary tract, esophagus, gastrointestinal system, liver, pancreas, fetus, placenta, chromaffin system, endocrine glands, ovary, uterus, penis, prostate, seminal vesicles, testis, bone marrow, immune system, cartilage, muscles, skeleton, central nervous system, ganglia, neuroglia, neurosecretory system, peripheral nervous system, bronchus, larynx, lung, nose, pleurus, ear, eye, mouth, pharynx, exocrine glands, bladder, kidney, ureter, and the like. The number of cDNA libraries selected can range from as few as 3 to greater than 10,000. Preferably, the number of the cDNA libraries is greater than 500.

In a preferred embodiment, gene sequences are assembled to reflect related sequences, such as assembled sequence fragments derived from a single transcript. Assembly of the polynucleotide sequences can be performed using sequences of various types including, but not limited to, ESTs, extensions, or shotgun sequences. In a most preferred embodiment, the polynucleotide sequences are derived from human sequences that have been assembled using the algorithm disclosed in "Database and System for Storing, Comparing and Displaying Related Biomolecular Sequence Information", Lincoln et al., Serial No:60/079,469, filed March 26, 1998, incorporated herein by reference.

Experimentally, differential expression of the polynucleotides can be evaluated by methods including, but not limited to, differential display by spatial immobilization or by gel electrophoresis, genome mismatch scanning, representational difference analysis, and transcript imaging. Additionally, differential expression can be assessed by microarray technology. These methods may be used alone or in combination.

Known corticosteroid synthesis genes can be selected based on the use of the genes as diagnostic or prognostic markers or as therapeutic targets for diseases associated with corticosteroid synthesis or steroid imbalance, more particularly, contraceptive disorders and infertility. Preferably, the known corticosteroid synthesis genes include steroid acute regulatory (StAR) gene, P450scc cholesterol side-chain cleavage enzyme (P450scc), 3-beta-hydroxysteroid dehydrogenase (3-beta-dehydrogenase), Type I 3-beta-hydroxysteroid dehydrogenase (Type I 3-beta-dehydrogenase), Type II 3-beta-hydroxysteroid dehydrogenase (Type II 3-beta-dehydrogenase), P450c11 beta-hydroxylase (11 beta-hydroxylase), and P450c17 alpha-hydroxylase (17-alpha-hydroxylase), and the like.

The procedure for identifying novel genes that exhibit a statistically significant coexpression pattern with known corticosteroid synthesis genes is as follows. First, the presence or absence of a gene

sequence in a cDNA library is defined: a gene is present in a cDNA library when at least one cDNA fragment corresponding to that gene is detected in a cDNA sample taken from the library, and a gene is absent from a library when no corresponding cDNA fragment is detected in the sample.

Second, the significance of gene coexpression is evaluated using a probability method to measure a due-to-chance probability of the coexpression. The probability method can be the Fisher exact test, the chi-squared test, or the kappa test. These tests and examples of their applications are well known in the art and can be found in standard statistics texts (Agresti (1990) Categorical Data Analysis, John Wiley & Sons. New York NY; Rice (1988) Mathematical Statistics and Data Analysis, Duxbury Press, Pacific Grove CA). A Bonferroni correction (Rice, supra, page 384) can also be applied in combination with one of the probability methods for correcting statistical results of one gene versus multiple other genes. In a preferred embodiment, the due-to-chance probability is measured by a Fisher exact test, and the threshold of the due-to-chance probability is set to less than 0.001, more preferably less than 0.00001.

To determine whether two genes, A and B, have similar coexpression patterns, occurrence data vectors can be generated as illustrated in Table 1, wherein a gene's presence is indicated by a one and its absence by a zero. A zero indicates that the gene did not occur in the library, and a one indicates that it occurred at least once.

Table 1. Occurrence data for genes A and B

	Library 1	Library 2	Library 3	...	Library N
gene A	1	1	0	...	0
gene B	1	0	1	...	0

For a given pair of genes, the occurrence data in Table 1 can be summarized in a 2 x 2 contingency table.

Table 2. Contingency table for co-occurrences of genes A and B

	Gene A present	Gene A absent	Total
Gene B present	8	2	10
Gene B absent	2	18	20
Total	10	20	30

Table 2 presents co-occurrence data for gene A and gene B in a total of 30 libraries. Both gene A and gene B occur 10 times in the libraries. Table 2 summarizes and presents: 1) the number of times gene A and B are both present in a library, 2) the number of times gene A and B are both absent in a library, 3) the number of times gene A is present while gene B is absent, and 4) the number of times gene B is present while gene A is absent. The upper left entry is the number of times the two genes co-occur in a

library, and the middle right entry is the number of times neither gene occurs in a library. The off diagonal entries are the number of times one gene occurs while the other does not. Both A and B are present eight times and absent 18 times, gene A is present while gene B is absent two times, and gene B is present while gene A is absent two times. The probability ("p-value") that the above association occurs due to chance as calculated using a Fisher exact test is 0.0003. Associations are generally considered significant if a p-value is less than 0.01 (Agresti, supra; Rice, supra).

This method of estimating the probability for coexpression of two genes makes several assumptions. The method assumes that the libraries are independent and are identically sampled. However, in practical situations, the selected cDNA libraries are not entirely independent because more than one library may be obtained from a single patient or tissue, and they are not entirely identically sampled because different numbers of cDNA's may have been sequenced from each library (typically ranging from 5,000 to 10,000 cDNA's per library). In addition, because a Fisher exact coexpression probability is calculated for each gene versus 41,419 other genes, a Bonferroni correction for multiple statistical tests is necessary.

Using the method of the present invention, we have identified seven novel genes that exhibit strong association, or coexpression, with known genes that are corticosteroid synthesis-specific. The known corticosteroid synthesis genes include steroid acute regulatory (StAR) gene, P450scc cholesterol side-chain cleavage enzyme (P450scc), 3-beta-hydroxysteroid dehydrogenase (3-beta-dehydrogenase), Type I 3-beta-hydroxysteroid dehydrogenase (Type I 3-beta-dehydrogenase), Type II 3-beta-hydroxysteroid dehydrogenase (Type II 3-beta-dehydrogenase), P450c11 beta-hydroxylase (11 beta-hydroxylase), and P450c17 alpha-hydroxylase (17-alpha-hydroxylase). The results presented in Table 5 show that the expression of the seven novel genes have direct or indirect association with the expression of known corticosteroid synthesis genes. Therefore, the novel genes can potentially be used in diagnosis, treatment, prognosis, or prevention of diseases associated with corticosteroid synthesis, or in the evaluation of therapies for diseases associated with corticosteroid synthesis. Further, the gene products of the seven novel genes are potential therapeutic proteins and targets of therapeutics against diseases associated with corticosteroid synthesis.

Therefore, in one embodiment, the present invention encompasses a polynucleotide sequence comprising the sequence of SEQ ID NOs:1-7. These seven polynucleotides are shown by the method of the present invention to have strong coexpression association with known corticosteroid synthesis genes and with each other. The invention also encompasses a variant of the polynucleotide sequence, its complement, or 18 consecutive nucleotides of a sequence provided in the above described sequences. Variant polynucleotide sequences typically have at least about 70%, more preferably at least about 85%, and most preferably at least about 95% polynucleotide sequence identity to NSEQ.

One preferred method for identifying variants entails using NSEQ and/or PSEQ sequences to search against the GenBank primate (pri), rodent (rod), mammalian (mam), vertebrate (vrtp), and eukaryote (eukp) databases, SwissProt, BLOCKS (Bairoch (1997) Nucleic Acids Res. 25:217-221), PFAM, and other databases that contain previously identified and annotated motifs, sequences, and gene functions. Methods that search for primary sequence patterns with secondary structure gap penalties (Smith (1992) Prot. Eng. 5:35-51) as well as algorithms such as BLAST (Basic Local Alignment Search Tool; Altschul (1993) J. Mol. Evol 36:290-300; and Altschul et al. (1990) J. Mol. Biol. 215:403-410), BLOCKS (Henikoff and Henikoff (1991) Nucleic Acids Research 19:6565-6572), Hidden Markov Models (HMM; Eddy (1996) Cur. Opin. Str. Biol. 6:361-365; and Sonnhammer et al. (1997) Proteins 28:405-420), and the like, can be used to manipulate and analyze nucleotide and amino acid sequences. These databases, algorithms and other methods are well known in the art and are described in Ausubel et al. (1997; Short Protocols in Molecular Biology, John Wiley & Sons, New York NY)and in Meyers (1995; Molecular Biology and Biotechnology, Wiley VCH, Inc, New York NY, p 856-853).

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to SEQ ID NOs:1-7, and fragments thereof under stringent conditions. Stringent conditions can be defined by salt concentration, temperature, and other chemicals and conditions well known in the art. In particular, stringency can be increased by reducing the concentration of salt, or raising the hybridization temperature.

For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and most preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Stringent temperature conditions will ordinarily include temperatures of at least about 30°C, more preferably of at least about 37°C, and most preferably of at least about 42°C. Varying additional parameters, such as hybridization time, the concentration of detergent or solvent, and the inclusion or exclusion of carrier DNA, are well known to those skilled in the art. Additional variations on these conditions will be readily apparent to those skilled in the art (Wahl and Berger (1987) Methods Enzymol. 152:399-407; Kimmel (1987) Methods Enzymol. 152:507-511; Ausubel, supra; and Sambrook et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY).

NSEQ or the polynucleotide sequences encoding PSEQ can be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. (See, e.g., Dieffenbach and Dveksler (1995) PCR Primer, a Laboratory Manual, Cold Spring Harbor Press, Plainview NY; Sarkar (1993) PCR Methods Applic. 2:318-322; Triglia et al. (1988) Nucleic Acids Res. 16:8186; Lagerstrom et al. (1991) PCR Methods Applic. 1:111-119; and Parker et al. (1991) Nucleic Acids Res. 19:3055-306.) Additionally, one

may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 18 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

In another aspect of the invention, NSEQ or the polynucleotide sequences encoding PSEQ can be cloned in recombinant DNA molecules that direct expression of PSEQ or the polypeptides encoded by NSEQ, or structural or functional fragments thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express the polypeptides of PSEQ or the polypeptides encoded by NSEQ. The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter the nucleotide sequences for a variety of purposes including, but not limited to, modification of cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

In order to express a biologically active polypeptide encoded by NSEQ, NSEQ, or the polynucleotide sequences encoding PSEQ, or derivatives thereof, may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions and NSEQ or polynucleotide sequences encoding PSEQ. Methods which are well known to those skilled in the art may be used to construct expression vectors containing NSEQ or polynucleotide sequences encoding PSEQ and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook (supra) and Ausubel, (supra).)

A variety of expression vector/host cell systems may be utilized to contain and express NSEQ or polynucleotide sequences encoding PSEQ. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (baculovirus); plant cell systems transformed with viral expression vectors, cauliflower mosaic virus

(CaMV) or tobacco mosaic virus (TMV), or with bacterial expression vectors (Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed. For long term production of recombinant proteins in mammalian systems, stable expression of a polypeptide encoded by NSEQ in cell lines is preferred. For example, NSEQ or sequences encoding PSEQ can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector.

In general, host cells that contain NSEQ and that express PSEQ may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences. Immunological methods for detecting and measuring the expression of PSEQ using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS).

Host cells transformed with NSEQ or polynucleotide sequences encoding PSEQ may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing NSEQ or polynucleotides encoding PSEQ may be designed to contain signal sequences which direct secretion of PSEQ or polypeptides encoded by NSEQ through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC, Manassas, MD) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant NSEQ or nucleic acid sequences encoding PSEQ are ligated to a heterologous sequence resulting in translation of a fusion protein containing heterologous protein moieties in any of the aforementioned host systems. Such heterologous protein moieties facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose

binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, hemagglutinin (HA) and monoclonal antibody epitopes..

In another embodiment, NSEQ or sequences encoding PSEQ are synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers (1980)Nucleic Acids Symp. Ser. (7):215-223; Horn et al. (1980) Nucleic Acids Symp. Ser. (7):225-232; and Ausubel, supra.) Alternatively, PSEQ or a polypeptide sequence encoded by NSEQ itself, or a fragment thereof, may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge et al. (1995) Science 269:202-204). Automated synthesis may be achieved using the ABI 431A Peptide synthesizer (PE Biosytems, Foster City CA). Additionally, PSEQ or the amino acid sequence encoded by NSEQ, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a polypeptide variant.

In another embodiment, the invention provides a substantially purified polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:9 or fragments thereof.

DIAGNOSTICS and THERAPEUTICS

The sequences of the these genes can be used in diagnosis, prognosis, treatment, prevention, and evaluation of therapies for diseases associated with corticosteroid synthesis particularly diseases associated with corticosteroid synthesis or steroid imbalance; and to the use of these biomolecules as therapeutics in reproductive medicine including contraception and fertility.

In one preferred embodiment, the polynucleotide sequences of NSEQ or the polynucleotides encoding PSEQ are used for diagnostic purposes to determine the absence, presence, and excess expression of PSEQ. The polynucleotides may be at least 10, preferably 18 nucleotides long, complementary RNA and DNA molecules, branched nucleic acids, and peptide nucleic acids (PNAs). The polynucleotides may be used to detect and quantitate gene expression in samples in which altered expression of PSEQ or the polypeptides encoded by NSEQ are correlated with disease. Alternatively, the polynucleotides may be used to monitor the levels of NSEQ or the polypeptides encoded by NSEQ during therapeutic intervention. Additionally, NSEQ or the polynucleotides encoding PSEQ can be used to detect genetic polymorphisms associated with a disease. These polymorphisms may be detected at the transcript cDNA or genomic level from mapping experiments.

The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low), will determine whether the probe identifies only naturally occurring sequences encoding PSEQ, allelic variants, or other related sequences.

Probes may also be used for the detection of related sequences, and should preferably have at least 70% sequence identity to any of the NSEQ or PSEQ-encoding sequences.

Means for producing specific hybridization probes for DNAs encoding PSEQ include the cloning of NSEQ or polynucleotide sequences encoding PSEQ into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, by fluorescent labels, and the like. The polynucleotide sequences encoding PSEQ may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; and in microarrays utilizing fluids or tissues from patients to detect altered PSEQ expression. Such qualitative or quantitative methods are well known in the art.

NSEQ or the nucleotide sequences encoding PSEQ can be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed, and the signal is quantitated and compared with a standard value, typically, derived from a non-diseased sample. If the amount of signal in the patient sample is altered in comparison to the standard value then the presence of altered levels of nucleotide sequences of NSEQ and those encoding PSEQ in the sample indicates the presence of the associated disease. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

Once the presence of a disease is established and a treatment protocol is initiated, hybridization or amplification assays can be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in a healthy subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

The polynucleotides may be used for the diagnosis of a variety of diseases associated with corticosteroid synthesis, particularly for cardiovascular disease, breast cancer, prostate cancer, osteoporosis, diabetes, and menopausal symptoms.

Alternatively, the polynucleotides may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify splice variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disease, to diagnose a disease, and to develop and monitor the activities of therapeutic agents.

In yet another alternative, polynucleotides may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence and detecting genetic diversity. Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich et al. (1995) In: Meyers (ed.) supra, pp. 965-968). Microarrays may be used to detect genetic diversity at the genome level.

In another embodiment, antibodies which specifically bind PSEQ may be used for the diagnosis of diseases characterized by the over-or-underexpression of PSEQ or polypeptides encoded by NSEQ. A variety of protocols for measuring PSEQ or the polypeptides encoded by NSEQ, including ELISAs, RIAs, and FACS, are well known in the art and provide a basis for diagnosing altered or abnormal levels of the expression of PSEQ or the polypeptides encoded by NSEQ. Standard values for PSEQ expression are established by combining body fluids or cell extracts taken from healthy subjects, preferably human, with antibody to PSEQ or a polypeptide encoded by NSEQ under conditions suitable for complex formation The amount of complex formation may be quantitated by various methods, preferably by photometric means. Quantities of PSEQ or the polypeptides encoded by NSEQ expressed in disease samples from, for example, biopsied tissues are compared with standard values. Deviation between standard and subject values establishes the parameters for diagnosing or monitoring disease. Alternatively, one may use competitive drug screening assays in which neutralizing antibodies capable of binding PSEQ or the polypeptides encoded by NSEQ specifically compete with a test compound for binding the polypeptides. Antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with PSEQ or the polypeptides encoded by NSEQ.

In another aspect, the polynucleotides and polypeptides of the present invention can be employed for treatment of diseases associated with the altered expression of novel corticosteroid synthesis-associated genes. The polynucleotides of NSEQ or those encoding PSEQ, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotides of NSEQ or those encoding PSEQ may be used in situations in which it would be desirable to block the transcription or translation of the mRNA using antisense technologies.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding PSEQ. (See, e.g., Sambrook, supra; and Ausubel, supra.)

Genes having polynucleotide sequences of NSEQ or those encoding PSEQ can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding PSEQ. Such constructs may be used to introduce untranslatable sense or

antisense sequences into a cell. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee et al. In: Huber and Carr, (1994) Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.)

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the cleavage of mRNA and decrease the levels of particular mRNAs, such as those comprising the polynucleotide sequences of the invention. (See, e.g. Rossi, 1994, Current Biology 4: 469-471.) Ribozymes may cleave mRNA at specific cleavage sites. Alternatively, ribozymes may cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The construction and production of ribozymes is well known in the art and is described in Meyers (supra).

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. Alternatively, nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases may be included.

Alternatively, the polynucleotides of the invention may be integrated into a genome by somatic or germ cell gene therapy. Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nature Biotechnology 15:462-466.)

Additionally, endogenous polynucleotide expression may be inactivated using homologous recombination methods which insert inactive gene sequence at the target sequence location. (See, e.g., Thomas and Capecchi (1987) Cell 51: 503-512.)

Further, an antagonist or antibody of a polypeptide of PSEQ or encoded by NSEQ may be administered to a subject to treat or prevent a cancer associated with increased expression or activity of PSEQ. An antibody which specifically binds the polypeptide may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express the the polypeptide.

Antibodies to PSEQ or polypeptides encoded by NSEQ may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use. Monoclonal antibodies to PSEQ may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. In addition, techniques developed for the production of chimeric antibodies can be used. (See, e.g., Meyers, supra.) Alternatively, techniques described for the production of single chain antibodies may be employed. Antibody fragments which contain specific binding sites for PSEQ or the polypeptide sequences encoded by NSEQ may also be generated.

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art.

Yet further, an agonist of a polypeptide of PSEQ or that encoded by NSEQ may be administered to a subject to treat or prevent a cancer associated with decreased expression or activity of the polypeptide.

An additional aspect of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of polypeptides of PSEQ or those encoded by NSEQ, antibodies to the polypeptides, and mimetics, agonists, antagonists, or inhibitors of the polypeptides. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's

Pharmaceutical Sciences (Maack Publishing, Easton PA).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example, polypeptides of PSEQ or those encoded by NSEQ, or fragments thereof, antibodies of the polypeptides, and agonists, antagonists or inhibitors of the polypeptides, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics.

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

INDUSTRIAL APPLICABILITY

It is understood that this invention is not limited to the particular methodology, protocols, and reagents described, as these may vary. It is also understood that the terminology used herein is for the purpose of describing particular embodiments only and is not intended to limit the scope of the present invention which will be limited only by the appended claims. The examples below are provided to illustrate the subject invention and are not included for the purpose of limiting the invention.

I cDNA Library Construction

The cDNA library, ADRENOT07, was selected to demonstrate the construction of the cDNA libraries from which the sequences used to identify genes associated with corticosteroid synthesis were derived. The ADRENOT07 cDNA library was constructed from microscopically normal adrenal tissues obtained from a 61-year old Caucasian female. Pathology indicated no significant abnormality of the right and left adrenals. Patient history included the diagnosis of unspecified disorder of adrenal glands, depressive disorder, benign hypertension, vocal cord paralysis, hemiplegia, subarachnoid hemorrhage, communicating hydrocephalus, and neoplasm of uncertain behavior of pituitary gland and craniopharyngeal duct. Prior surgery included total excision of the pituitary gland. Family history included malignant prostate neoplasm in the father and malignant colon neoplasm in the mother.

The frozen tissue was homogenized and lysed using a POLYTRON homogenizer (PT-3000; Brinkmann Instruments, Westbury NY) in guanidinium isothiocyanate solution. The lysate was centrifuged over a 5.7 M CsCl cushion using an SW28 rotor in a BL8-70M ultracentrifuge (Beckman

Coulter, Fullerton CA) for 18 hours at 25,000 rpm at ambient temperature. The RNA was extracted with acid phenol, pH 4.7, precipitated using 0.3 M sodium acetate and 2.5 volumes of ethanol, resuspended in RNAse-free water, and treated with DNase at 37°C. The RNA extraction was repeated with acid phenol, pH 4.7, and precipitated with sodium acetate and ethanol as before. The mRNA was tisolated using the OLIGOTEX kit (Qiagen, Chatsworth CA) and used to construct the cDNA library.

The mRNA was handled according to the recommended protocols in the SUPERSCRIPT Plasmid system (Life Technologies, Gaithersburg MD). The cDNAs were fractionated on a SEPHAROSE CL4B column (Amersham Pharmacia Biotech, Piscataway NJ), and those cDNAs exceeding 400 bp were ligated into pINCY 1 plasmid (Incyte Pharmaceuticals, Palo Alto CA). The plasmid was subsequently transformed into DH5α competent cells (Life Technologies).

II Isolation and Sequencing of cDNA Clones

Plasmid DNA was released from the cells and purified using the REAL Prep 96 Plasmid kit (Qiagen). This kit enabled the simultaneous purification of 96 samples in a 96-well block using multi-channel reagent dispensers. The recommended protocol was employed except for the following changes: 1) the bacteria were cultured in 1 ml of sterile Terrific Broth (Life Technologies) with carbenicillin at 25 mg/L and glycerol at 0.4%; 2) after inoculation, the cultures were incubated for 19 hours and at the end of incubation, the cells were lysed with 0.3 ml of lysis buffer; and 3) following isopropanol precipitation, the plasmid DNA pellet was resuspended in 0.1 ml of distilled water. After the last step in the protocol, samples were transferred to a 96-well block for storage at 4° C.

The cDNAs were prepared using a MICROLAB 2200 (Hamilton, Reno, NV) in combination with DNA ENGINE thermal cyclers (PTC200; MJ Research, Watertown, MA) and sequenced by the method of Sanger et al. (1975, J. Mol. Biol. 94:441f) using ABI PRISM 377 DNA Sequencing systems (PE Biosystems).

III Selection, Assembly, and Characterization of Sequences

The sequences used for coexpression analysis were assembled from EST sequences, 5' and 3' longread sequences, and full length coding sequences. Selected assembled sequences were expressed in at least three cDNA libraries.

The assembly process is described as follows. EST sequence chromatograms were processed and verified. Quality scores were obtained using PHRED (Ewing et al. (1998) Genome Res. 8:175-185; Ewing and Green (1998) Genome Res. 8:186-194). Then the edited sequences were loaded into a relational database management system (RDBMS). The EST sequences were clustered into an initial set of bins using BLAST with a product score of 50. All clusters of two or more sequences were created as bins. The overlapping sequences represented in a bin correspond to the sequence of a transcribed gene.

Assembly of the component sequences within each bin was performed using a modification of

PHRAP, a publicly available program for assembling DNA fragments (Green, University of Washington, Seattle WA). Bins that showed 82% identity from a local pair-wise alignment between any of the consensus sequences were merged.

Bins were annotated by screening the consensus sequence in each bin against public databases, such as GBpri and GenPept from NCBI. The annotation process involved a FASTn screen against the GBpri database in GenBank. Those hits with a percent identity of greater than or equal to 70% and an alignment length of greater than or equal to 100 base pairs were recorded as homolog hits. The residual unannotated sequences were screened by FASTx against GenPept. Those hits with an E value of less than or equal to 10^{-8} are recorded as homolog hits.

Sequences were then reclustered using BLASTn and Cross-Match, a program for rapid protein and nucleic acid sequence comparison and database search (Green, supra), sequentially. Any BLAST alignment between a sequence and a consensus sequence with a score greater than 150 was realigned using cross-match. The sequence was added to the bin whose consensus sequence gave the highest Smith-Waterman score amongst local alignments with at least 82% identity. Non-matching sequences created new bins. The assembly and consensus generation processes were performed for the new bins.

IV Coexpression Analyses of Known Corticosteroid Synthesis Genes

Seven known corticosteroid synthesis genes were selected to identify novel genes that are closely associated with corticosteroid synthesis. These known genes were steroid acute regulatory (StAR) gene, P450scc cholesterol side-chain cleavage enzyme (P450scc), 3-beta-hydroxysteroid dehydrogenase (3-beta-dehydrogenase), Type I 3-beta-hydroxysteroid dehydrogenase (Type I 3-beta-dehydrogenase), Type II 3-beta-hydroxysteroid dehydrogenase (Type II 3-beta-dehydrogenase), P450c11 beta-hydroxylase (11 beta-hydroxylase), and P450c17 alpha-hydroxylase (17-alpha-hydroxylase). Corticosteroid synthesis occurs primarily in the adrenal cortex. The proteins encoded by the corticosteroid synthesis genes examined here are six enzymes that catalyze steroid synthesis and one protein that transports cholesterol, the starting substrate, to the locus of the first enzyme in the pathway and thereby initiates synthesis. The principal substrates and products of these enzymes are cholesterol, progesterone, pregnenolone, hydroxypregnenolone, hydroxyprogesterone, corticosterone and aldosterone. These products are modified further (in the testes or ovaries) to produce testosterone, estrogens, and other steroids.

The known corticosteroid synthesis genes, that we examined in this analysis, and brief descriptions of their functions are listed in Table 4. Detailed descriptions of their roles in corticosteroid synthesis may be found in the cited articles and reviews.

Table 4. Known corticosteroid-synthesis genes.

Gene	Description & references
StAR	Steroid acute regulatory (StAR) gene
	Transports cholesterol from the cytoplasm into the mitochondria, locus of the first steps in steroid synthesis (Gradi et al. (1995) Biochim Biophys Acta 1258: 228-33; Norman & Litwack, supra)
P450scc	P450scc cholesterol side-chain cleavage enzyme
	Catalyzes 20R-hydroxylation, 22R-hydroxy-lation, and cholesterol scission of C-20-C-22 carbon bonds; converts cholesterol to pregnenolone (Laycock & Wise, supra; Norman & Litwack, supra)
3-beta-dehydrogenase	3-beta-hydroxysteroid dehydrogenase
	Catalyzes 5-ene 3-beta-hydroxysteroid to 4-ene 3-oxosteroid, converts pregnenolone to progesterone, converts 17OH-pregnenolone to 17OH-progesterone; converts dehydroepian-drosterone to androstenedione; occurs in multiple forms ((Laycock & Wise, supra; Lorence et al. (1990) Endocrinology 126: 2493-8; Norman & Litwack, supra)
Type I 3-beta dehydrogenase	Variant form of 3-beta-hydroxysteroid de-hydrogenase, or 3-beta dehydrogenase (Rheaume et al. (1991) Mol Endocrinol 5: 1147-57)
Type II 3-beta dehydrogenase	Variant form of 3-beta-hydroxysteroid dehydrogenas (Rheaume et al., supra)
11 beta-hydroxylase	P450c11 beta-hydroxylase
	Catalyzes 11beta hydroxylation, 18-hydroxylation, and oxidation or C-18 hydroxyl to aldehyde; converts deoxycortisol to cortisol; converts deoxycorticosterone to corticosterone; converts corticosterone to aldosterone (Laycock & Wise, supra; Norman & Litwack, supra)
17-alpha-hydroxylase	P450c17 alpha-hydroxylase
	Catalyzes 17-alpha hydroxylation and C-17-C-20 scission; converts pregnenolone to 17OH-pregnenolone; converts 17OH-pregnenolone to dehydroepiandrosterone; onverts 17OH-progesterone to androstenedione (Laycock & Wise, supra; Norman & Litwack, , supra)

The coexpression of the seven known genes with each other is shown in Table 5. The entries in Table 5 are the negative log of the p-value (- log p) for the coexpression of the two genes. As shown, the method successfully identified the strong association of the known genes among themselves, indicating that the coexpression analysis method of the present invention was effective in identifying genes that are

closely associated with corticosteroid synthesis.

Table 5. Co-expression of 7 novel genes and 7 known steroid-synthesis genes. (- log *p*)

	3-beta-dehydrogenase	11-beta-	17-alpha-hydroxylase	StAR	P450scc	Type I 3-beta-hydroxylase	3-beta-dehydrogenase	2737624	2364582	64973	2961563	65781	1419725	2867065
3-beta-dehydrogenase														
11-beta-hycroxylase	10													
17-alpha-hydroxylase	10	1												
StAR	8	1	15											
P450scc	14	8	9	8										
Type I 3-beta-hydroxylase	8	7	7	9	7									
3-beta-dehydrogenase	15	4	5	3	1	5								
2737624	8	9	7	7	7	8	5							
2364582	4	5	4	4	4	3	3	6						
64973	6	4	3	3	3	3	2	3	0					
2961563	5	8	8	5	6	4	3	4	2	2				
65781	7	0	1	0	8	0	7	0	0	1	2			
1419725	3	5	4	5	5	4	1	4	3	1	6	0		
2867065	3	6	4	4	3	3	1	4	3	2	4	0	4	

V Novel Genes Associated with Corticosteroid Synthesis

Using coexpression analysis, we have identified seven novel genes that show strong association with known corticosteroid synthesis genes from a total of 41,419 assembled gene sequences. The degree of association was measured by probability values and has a cutoff of p value less than 0.00001. Identification was followed by annotation and literature searches to insure that the genes that passed the probability test have strong association with known corticosteroid synthesis genes. This process was reiterated so that the initial 41,419 genes were reduced to the final seven corticosteroid synthesis-associated genes. Details of the expression patterns for the seven novel corticosteroid synthesis genes were presented above in Table 5.

Each of the seven novel genes is coexpressed with at least one of the seven known genes with a p-value of less than $10E^{-05}$. The coexpression results are shown in Table 5. The novel genes identified are listed in the table by their Incyte clone numbers (Clone), and the known genes by their abbreviated names (Gene) as shown in Example IV.

VI Novel Genes Associated with Corticosteroid Synthesis

Seven novel genes were identified from the data shown in Table 5 to be associated with corticosteroid synthesis.

Nucleic acids comprising the consensus sequences of SEQ ID NOs:1-7 of the present invention

were first identified from Incyte Clones 64973, 65781, 1419725, 2364582, 2737624, 2867065, and 2961563, respectively, and assembled according to Example III. BLAST and other motif searches were performed for SEQ ID NOs:1-7 according to Example VII. The sequences of SEQ ID NOs:1-7 were translated and sequence identity was sought with known sequences. Amino acids comprising the consensus sequences of SEQ ID NO:8 and SEQ ID NO:9 of the present invention were encoded by the nucleic acids of SEQ ID NO:2 and SEQ ID NO:6, respectively. SEQ ID NOs:8 and 9 were also analyzed using BLAST and other motif search tools as disclosed in Example VII.

SEQ ID NO:4 is 567 nucleotides in length and shows about 68% sequence identity from about nucleotide 205 to about nucleotide 507 with human mRNA for alpha 1C adrenergic receptor isoform 2 (g927208). SEQ ID NO:5 is 920 nucleotides in length and shows about 78% sequence identity from about nucleotide 649 to about nucleotide 920 and from about nucleotide 8 to about nucleotide 153 with a human glucose phosphate isomerase mRNA (g309269). Glucose phosphate isomerase is a housekeeping gene expressed in all tissues and organisms that utilize glycolysis and gluconeogenesis.

SEQ ID NO:8 is 334 amino acid residues in length and shows about 92% sequence identity from about amino acid residue 1 to about amino acid residue 181 with KIAA0686 (g3327186), a human protein encoded by a gene from the brain tissue. The sequence encompassing residues 1 to 22 of SEQ ID NO:8 is a potential signal peptide, as shown by SPSCAN analysis according to Example VII. SEQ ID NO:8 also has one potential casein kinase II phosphorylation site at S169; one potential N-myristoylation site at G60; and four potential protein kinase C phosphorylation sites at T74, S101, S129, and S156. SEQ ID NO:9 is 334 amino acid residues in length and shows about 99% sequence identity from about amino acid residue 154 to about amino acid residue 257 with a secreted protein encoded by clone AS162_1 (WO 97/46683). The sequence encompassing residues 1-50 of SEQ ID NO:9 is a potential signal peptide by SPSCAN analysis according to Example VII. HMM analysis shows that SEQ ID NO:9 has four potential transmembrane domains encompassing amino acid residues 38 to 60, 89 to 106, 135 to 152, and 160 to 185. SEQ ID NO:9 also has one potential N-glycosylation site at N304; three potential casein kinase II phosphorylation sites at S300, S302, and T315; and one potential protein kinase C phosphorylation site at T34.

VII Homology Searching for Corticosteroid Synthesis Genes and the Proteins

Polynucleotide sequences, SEQ ID NOs:1-7, and polypeptide sequences, SEQ ID NOs:8 and 9, were queried against databases derived from sources such as GenBank and SwissProt. These databases, which contain previously identified and annotated sequences, were searched for regions of similarity using Basic Local Alignment Search Tool (BLAST; Altschul (1990, supra) and Smith-Waterman alignment (Smith, supra). BLAST searched for matches and reported only those that satisfied the

probability thresholds of 10^{-25} or less for nucleotide sequences and 10^{-8} or less for polypeptide sequences.

The polypeptide sequences were also analyzed for known motif patterns using MOTIFS, SPSCAN, BLIMPS, and Hidden Markov Model (HMM)-based protocols. MOTIFS (Genetics Computer Group, Madison WI) searches polypeptide sequences for patterns that match those defined in the Prosite Dictionary of Protein Sites and Patterns (Bairoch, supra), and displays the patterns found and their corresponding literature abstracts. SPSCAN (Genetics Computer Group) searches for potential signal peptide sequences using a weighted matrix method (Nielsen et al. (1997) Prot. Eng. 10: 1-6). Hits with a score of 5 or greater were considered. BLIMPS uses a weighted matrix analysis algorithm to search for sequence similarity between the polypeptide sequences and those contained in BLOCKS, a database consisting of short amino acid segments, or blocks, of 3-60 amino acids in length, compiled from the PROSITE database (Henikoff and Henikoff, supra; Bairoch et al. supra), and those in PRINTS, a protein fingerprint database based on non-redundant sequences obtained from sources such as SwissProt, GenBank, PIR, and NRL-3D (Attwood et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424). For the purposes of the present invention, the BLIMPS searches reported matches with a cutoff score of 1000 or greater and a cutoff probability value of 1.0×10^{-3}. HMM-based protocols were based on a probabilistic approach and searched for consensus primary structures of gene families in the protein sequences (Eddy, supra; Sonnhammer et al. supra). More than 500 known protein families with cutoff scores ranging from 10 to 50 bits were selected for use in this invention.

VIII Labeling and Use of Individual Hybridization Probes

Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (NEN Life Science Products, Boston, MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine resin column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba 1, or Pvu II (NEN Life Science Products).

The DNA from each digest is fractionated on a 0.7 percent agarose gel and transferred to NYTRANPLUS membranes (Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. After XOMAT AR film (Eastman Kodak, Rochester NY) is exposed to the blots for several hours, hybridization patterns are compared.

IX. Production of Specific Antibodies

SEQ ID NO:8 or 9, substantially purified using polyacrylamide gel electrophoresis (Harrington

(1990) Methods Enzymol. 182:488-495) or other purification techniques is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the amino acid sequence is analyzed using LASERGENE software (DNASTAR, Madison WI) to determine regions of high immunogenicity, and an oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. Typically, oligopeptides 15 residues in length are synthesized using an ABI 431A Peptide synthesizer (PE Biosystems) using Fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidoben-zoyl-N-hydroxysuccinimide ester to increase immunogenicity. Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide activity by, for example, binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

What is claimed is:

1. A substantially purified polynucleotide comprising a gene that is coexpressed with one or more known corticosteroid synthesis genes in a plurality of biological samples, wherein each known corticosteroid synthesis gene is selected from the group consisting of steroid acute regulatory gene, P450scc cholesterol side-chain cleavage enzyme, 3-beta-hydroxysteroid dehydrogenase, Type I 3-beta-hydroxysteroid dehydrogenase, Type II 3-beta-hydroxysteroid dehydrogenase, P450c11 beta-hydroxylase, and P450c17 alpha-hydroxylase.

2. The polynucleotide of claim 1, comprising a polynucleotide sequence selected from:

(a) a polynucleotide sequence selected from the group consisting of SEQ ID NOs:1- 7 ;

(b) a polynucleotide sequence which encodes the polypeptide sequence of SEQ ID NO:8 or 9;

(c) a polynucleotide sequence having at least 70% identity to the polynucleotide sequence of (a) or (b);

(d) a polynucleotide sequence comprising at least 18 sequential nucleotides of the polynucleotide sequence of (a), (b), or (c);

(e) a polynucleotide sequence which is complementary to the polynucleotide sequence of (a), (b),(c),(d) or (d); and

(f) a polynucleotide which hybridizes under stringent conditions to the polynucleotide of (a),(b), (c), (d) or (e).

3. A substantially purified polypeptide comprising the gene product of a gene that is coexpressed with one or more known corticosteroid synthesis genes in a plurality of biological samples, wherein each known corticosteroid synthesis gene is selected from the group consisting of steroid acute regulatory gene, P450scc cholesterol side-chain cleavage enzyme, 3-beta-hydroxysteroid dehydrogenase, Type I 3-beta-hydroxysteroid dehydrogenase, Type II 3-beta-hydroxysteroid dehydrogenase, P450c11 beta-hydroxylase, and P450c17 alpha-hydroxylase.

4. The polypeptide of claim 3, comprising a polypeptide sequence selected from:

(a) the polypeptide sequence of SEQ ID NO:8 or 9;

(b) a polypeptide sequence having at least 85% identity to the polypeptide sequence of (a); and

(c) a polypeptide sequence comprising at least 6 sequential amino acids of the polypeptide sequence of (a) or (b).

5. An expression vector comprising the polynucleotide of claim 2.

6. A host cell comprising the expression vector of claim 5.

7. A pharmaceutical composition comprising the polynucleotide of claim 2 or the polypeptide of claim 3 in conjunction with a suitable pharmaceutical carrier.

8. An antibody which specifically binds to the polypeptide of claim 4.

9. A method for diagnosing a disease or condition associated with the altered expression of a gene that is coexpressed with one or more known corticosteroid synthesis genes, wherein each known corticosteroid synthesis gene is selected from the group consisting of steroid acute regulatory gene, P450scc cholesterol side-chain cleavage enzyme, 3-beta-hydroxysteroid dehydrogenase, Type I 3-beta-hydroxysteroid dehydrogenase, Type II 3-beta-hydroxysteroid dehydrogenase, P450c11 beta-hydroxylase, and P450c17 alpha-hydroxylase, the method comprising the steps of:

(a) providing a sample comprising one of more of said coexpressed genes;

(b) hybridizing the polynucleotide of claim 2 to said coexpressed genes under conditions effective to form one or more hybridization complexes;

(c) detecting the hybridization complexes; and

(d) comparing the levels of the hybridization complexes with the level of hybridization complexes in a non-diseased sample, wherein altered expression levels correlate with the presence of the disease or condition.

10. A method for treating or preventing a disease associated with the altered expression of a gene that is coexpressed with one or more known corticosteroid synthesis genes in a subject in need, wherein each known corticosteroid synthesis gene is selected from the group consisting of steroid acute regulatory gene, P450scc cholesterol side-chain cleavage enzyme, 3-beta-hydroxysteroid dehydrogenase, Type I 3-beta-hydroxysteroid dehydrogenase, Type II 3-beta-hydroxysteroid dehydrogenase, P450c11 beta-hydroxylase, and P450c17 alpha-hydroxylase, the method comprising the step of administering to said subject in need the pharmaceutical composition of claim 7 in an amount effective for treating or preventing said disease.

11. A method for treating or preventing a disease associated with the altered expression of a gene that is coexpressed with one or more known corticosteroid synthesis genes in a subject in need, wherein each known corticosteroid synthesis gene is selected from the group consisting of steroid acute regulatory gene, P450scc cholesterol side-chain cleavage enzyme, 3-beta-hydroxysteroid dehydrogenase, Type I 3-beta-hydroxysteroid dehydrogenase, Type II 3-beta-hydroxysteroid dehydrogenase, P450c11 beta-hydroxylase, and P450c17 alpha-hydroxylase, the method comprising the step of administering to said subject in need the antibody of claim 8 in an amount effective for treating or preventing said disease.

12. A method for treating or preventing a disease associated with the altered expression of a gene that is coexpressed with one or more known corticosteroid synthesis genes in a subject in need, wherein each known corticosteroid synthesis gene is selected from the group consisting of steroid acute regulatory gene, P450scc cholesterol side-chain cleavage enzyme, 3-beta-hydroxysteroid dehydrogenase, Type I 3-beta-hydroxysteroid dehydrogenase, Type II 3-beta-hydroxysteroid dehydrogenase, P450c11 beta-hydroxylase, and P450c17 alpha-hydroxylase, the method comprising the step of administering to said subject in need the polynucleotide sequence of claim 2 in an amount effective for treating or preventing

said disease.

13. A ribozyme for cleaving the polynucleotide of claim 2.

14. A method for treating or preventing a disease or condition associated with the increased expression of a gene that is is coexpressed with one or more known corticosteroid synthesis genes in a subject in need, wherein each known corticosteroid synthesis gene is selected from the group consisting of steroid acute regulatory gene, P450scc cholesterol side-chain cleavage enzyme, 3-beta-hydroxysteroid dehydrogenase, Type I 3-beta-hydroxysteroid dehydrogenase, Type II 3-beta-hydroxysteroid dehydrogenase, P450c11 beta-hydroxylase, and P450c17 alpha-hydroxylase, the method comprising administering to a subject in need the ribozyme of claim 13 in an amount effective for treating or preventing said disease.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 May 2000 (18.05.2000)

PCT

(10) International Publication Number
WO 00/28027 A3

(51) International Patent Classification[7]: C12N 15/12, C07K 14/47, C12N 15/11, C07K 16/18, A61K 38/17

(21) International Application Number: PCT/US99/25457

(22) International Filing Date: 28 October 1999 (28.10.1999)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/155,269 6 November 1998 (06.11.1998) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 60/155,269 (CIP)
Filed on 6 November 1998 (06.11.1998)

(71) Applicant *(for all designated States except US)*: **INCYTE PHARMACEUTICALS, INC.** [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **WALKER, Michael, G.** [CA/US]; Unit 80, 1050 Borregas Avenue, Sunnyvale, CA 94089 (US). **VOLKMUTH, Wayne** [US/US]; 783 Roble Avenue #1, Menlo Park, CA 94025 (US). **KLINGLER, Tod, M.** [US/US]; 28 Dover Court, San Carlos, CA 94070 (US).

(74) Agents: **BILLINGS, Lucy, J.** et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States *(national)*: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *With international search report.*
— *Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.*

(88) Date of publication of the international search report:
30 November 2000

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 00/28027 A3

(54) Title: CORTICOSTEROID SYNTHESIS-ASSOCIATED GENES

(57) Abstract: The invention provides novel corticosteroid synthesis-associated genes and polypeptides encoded by those genes. The invention also provides expression vectors, host cells, antibodies, antisense molecules and ribozymes. The invention also provides methods for diagnosing, treating or preventing diseases associated with the altered expression of corticosteroid synthesis-associated genes.

Docket No.: PB-0005 USN

DECLARATION AND POWER OF ATTORNEY FOR UNITED STATES PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name, and

I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if more than one name is listed below) of the subject matter which is claimed and for which a United States patent is sought on the invention entitled

CORTICOSTEROID SYNTHESIS-ASSOCIATED GENES

the specification of which:

/ / is attached hereto.

/ X / was filed on May 1, 2001 as application Serial No. 09/830,913 and if this box contains an X/ /, was amended on _____.

/ / was filed as Patent Cooperation Treaty international application No. _____ on _____,2001, if this box contains an X /_/, was amended on under Patent Cooperation Treaty Article 19 on _____ 2001, and if this box contains an X /_/, was amended on _____

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge my duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim the benefit under Title 35, United States Code, §119 or §365(a)-(b) of any foreign application(s) for patent or inventor's certificate indicated below and of any Patent Cooperation Treaty international applications(s) designating at least one country other than the United States indicated below and have also identified below any foreign application(s) for patent or inventor's certificate and Patent Cooperation Treaty international application(s) designating at least one country other than the United States for the same subject matter and having a filing date before that of the application for said subject matter the priority of which is claimed:

Docket No.: PB-0005 USN

Country	Number	Filing Date	Priority Claimed
_____	_____	_____	/_/ Yes /_/ No
_____	_____	_____	/_/ Yes /_/ No

I hereby claim the benefit under Title 35, United States Code, §119(e) of any United States provisional application(s) listed below.

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)
60/155,269	November 6, 1998	Expired

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in said prior application(s) in the manner required by the first paragraph of Title 35, United States Code §112, I acknowledge my duty to disclose material information as defined in Title 37 Code of Federal Regulations, §1.56(a) which occurred between the filing date(s) of the prior application(s) and the national or Patent Cooperation Treaty international filing date of this application:

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)
PCT/US99/25457	October 28, 1999	Patented

I hereby appoint the following:

Lucy J. Billings	Reg. No. 36,749
Michael C. Cerrone	Reg. No. 39,132
Diana Hamlet-Cox	Reg. No. 33,302
Richard C. Ekstrom	Reg. No. 37,027
Barrie D. Greene	Reg. No. 46,740
Lynn E. Murry	Reg. No. 42,918
Shirley A. Recipon	Reg. No. 47,016
Cathleen M. Rocco	Reg. No. 46,172
Susan K. Sather	Reg. No. 44,316
Michelle M. Stempien	Reg. No. 41,327
David G. Streeter	Reg. No. 43,168

respectively and individually, as my patent attorneys and/or agents, with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith. Please address all communications to:

Docket No.: PB-0005 USN

LEGAL DEPARTMENT
INCYTE GENOMICS, INC.
3160 PORTER DRIVE, PALO ALTO, CA 94304

TEL: 650-855-0555 FAX: 650-849-8886 or 650-845-4166

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Sole Inventor or First Joint Inventor:

1 00

Full name: Michael G. Walker

Signature:

Date: 25 March 2002 , ~~2001~~

Citizenship United States of America

Residence: Sunnyvale, California CA

P.O. Address: 1050 Borregas Avenue, #80
Sunnyvale, California 94089

Second Joint Inventor:

2 00

Full name: Wayne Volkmuth

Signature:

Date: 18 , 2001

Citizenship United States of America

Residence: Calabasas, California CA

P.O. Address: 3620 El Encanto Dr.
Calabasas, California 91302

Docket No.: PB-0005 USN

Third Joint Inventor:

3rd

Full name:	**Tod Klingler**
Signature:	Tod Klingler
Date:	9 NOVEMBER **, 2001**
Citizenship	**United States of America**
Residence:	**San Carlos, California** CA
P.O. Address:	**28 Dover Court** **San Carlos, California 94070**

SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.
WALKER, Michael G.
VOLKMUTH, Wayne
KLINGLER, Tod M.

<120> CORTICOSTEROID SYNTHESIS-ASSOCIATED GENES

<130> PB-0005 PCT

<140> To Be Assigned
<141> Herewith

<150> 60/155,269
<151> 1998-11-06

<160> 9

<170> PERL Program

<210> 1
<211> 3796
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 064973CB1

<400> 1
tctttccagt tagagaagag gagtggtgat aaagcccact gaccttcaca cattctaaaa 60
attggttggc aatgccagta taccaacaat catgatcagc tgaaagaaac agatatttta 120
aattgccaga acaacaaatg atgaggcaac tacagtcaga tttatagcca gccatctatc 180
acctctagaa ggttccagag acagtgaaac tgcaagatgc tctcaacagg attatgtctc 240
atggagacca gtaagaaaat catttatctg aaggtgaaat gcagagttgg ataagaaata 300
cattgctggg tttctaaaat gctgccttcc tgcctctact ccacctccat ccctggactt 360
tggacccttg gcctaggagc ctaaggacct tcacccctgt gcaccaccca agaaagagga 420
aaactttgcc tacaactttg gaaatgctgg ggtccctggt gtggtaagaa actcaacatc 480
agacgggtat gcagaaggat gttcttctgg gatttgcagg tacataaaaa atgtatggca 540
tcttttcctt gcaaattctt ccagtttcca agtgagaagg ggagcaggtg tttactgatg 600
gaaaaggtat gttgctatgt tgatgtgtaa gtgaaatcag ttgtgtgcaa tagacagggg 660
cgtattcatg ggagcatcag ccagtttcta aaacccacag gccatcagca gctagaggtg 720
gctggctttg gccagacatg gaccctaaat caacagacaa tggcattgtc gaagagcaac 780
ctgttaatga atcatgttaa aaatcaaggt ttggcttcag tttaaatcac ttgaggtatg 840
aagtttatcc tgttttccag agataaacat aagttgatct tcccaaaata ccatcattag 900
gacctatcac acaatatcac tagtttttt tgtttgtttg ttttttgttt tgtttcttgg 960
taaagccatg caccacagac ttctgggcag agctgagaga caatggtcct gacataataa 1020
ggatctttga ttaaccccca taaggcatgt gtgtgtatac aaatatactt ctctttggct 1080
tttcgacata gaacctcagc tgttaaccaa ggggaaatac atcagatctg caacacagaa 1140
atgctctgcc tgaaatttcc accatgccta ggactcaccc catttatcca ggtctttctg 1200
gatctgttta atcaataagc cctataatca cttgctaaac actgggcttc atcacccagg 1260
gataaaaaca gagatcattg tcttggacct cctgcatcag cctattcaaa attatctctc 1320
tctctagctt tccacaaatc ctaaaattcc tgtcccaagc cacccaaatt ctcagatctt 1380
ttctggaaca aggcagaata taaaataaat atacatttag tggcttgggc tatggtctcc 1440
aaagatcctt caaaaataca tcaagccagc ttcattcact cactttactt agaacagaga 1500
tataagggcc tgggatgcat ttatttatc aataccaatt tttgtggcca tggcagacat 1560
tgctaatcaa tcacagcact atttcctatt aagcccactg atttcttcac aatccttctc 1620
aaattacaat tccaaagagc cgccactcaa cagtcagatg aacccaacag tcagatgaga 1680
gaaatgaacc ctacttgcta tctctatctt agaaagcaaa aacaaacagg agtttccagg 1740
gagaatggga aagccagggg gcataaaagg tacagtcagg ggaaaataga tctaggcaga 1800
gtgccttagt cagggaccac gggcgctgaa tctgcagtgc caacaccaaa ctgacacatc 1860
tccaggtgta cctccaaccc tagccttctc ccacagctgc ctacaacaga gtctcccagc 1920
cttctcagag agctaaaacc agaaatttcc agactcatga aagcaacccc ccagcctctc 1980
cccaaccctg ccgcattgtc taatttttag aacactaggc ttcttctttc atgtagttcc 2040

tcataagcag gggccagaat atctcagcca cctgcagtga cattgctgga cccctgaaaa 2100
ccattccata ggagaatggg ttccccaggc tcacagtgta gagacattga gcccatcaca 2160
actgttttga ctgctggcag tctaaaacag tccacccacc ccatggcact gccgcgtgat 2220
tcccgcggcc attcagaagt tcaagccgag atgctgacgt tgctgagcaa cgagatggtg 2280
agcatcagtg caaatgcacc attcagcaca tcagtcatat gcccagtgca gttacaagat 2340
gttgtttcgg caaagcattt tgatggaata gggaactgca aatgtatgat gattttgaaa 2400
aggctcagca ggatttgttc ttaaaccgac tcagtgtgtc atccccggtt atttagaatt 2460
acagttaaga aggagaaact tctataagac tgtatgaaca aggtgatatc ttcatagtgg 2520
gctattacag gcaggaaaat gttttaactg gtttacaaaa tccatcaata cttgtgtcat 2580
tccctgtaaa aggcaggaga catgtgatta tgatcaggaa actgcacaaa attattgttt 2640
tcagcccccg tgttattgtc cttttgaact gtttttttt ttattaaagc caaatttgtg 2700
ttgtatatat tcgtattcca tgtgttagat ggaagcattt cctatccagt gtgaataaaa 2760
agaacagttg tagtaaatta ttataaagcc gatgatattt catggcaggt tattctacca 2820
agctgtgctt gttggttttt cccatgactg tattgctttt ataaatgtac aaatagttac 2880
tgaaatgacg agacccttgt ttgcacagca ttaataagaa ccttgataag aaccatattc 2940
tgttgacagc cagctcacag tttcttgcct gaagcttggt gcaccctcca gtgagacaca 3000
agatctctct tttaccaaag ttgagaacag agctggtgga ttaattaata gtcttcgata 3060
tctggccatg ggtaacctca ttgtaactat catcagaatg ggcagagatg atcttgaagt 3120
gtcacataca ctaaagtcca aacactatgt cagatggggg taaaatccat taaagaacag 3180
gaaaaaataa ttataagatg ataagcaaat gtttcagccc aatgtcaacc cagttaaaaa 3240
aaaaattaat gctgtgtaaa atggttgaat tagtttgcaa actatataaa gacatatgca 3300
gtaaaaagtc tgttaatgca catcctgtgg gaatggagtg ttctaaccaa ttgccttttc 3360
ttgttatctg agctctccta tattatcata ctcagataac caaattaaaa gaattagaat 3420
atgattttta atacacttaa cattaaactc ttctaacttt cttctttctg tgataattca 3480
gaagatagtt atggatcttc aatgcctctg agtcattgtt ataaaaaatc agttatcact 3540
ataccatgct ataggagact gggcaaaacc tgtacaatga caaccctgga agttgctttt 3600
tttaaaaaaa taataaattt cttaaatcaa ctctttttc tggttgtctg tttgttataa 3660
agtgcaacgt attcaagtcc tcaatatcct gatcataata ccatgctata ggagactggg 3720
caaaacctgt acaatgacaa ccctggaagt tgctttttta aaaaaaataa taaatttctt 3780
aaatcaaaaa aaaaaa 3796

<210> 2
<211> 1187
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 065781CB1

<400> 2
atatacgatt actataggga tttggccctc gaggcaagaa attcggcacg aggatcagac 60
catcagaagg atttgtataa agagtgactc tcctatgaag gtaaaggcca cccctcttca 120
gttccagtga ctgagataca tttttccaat cctgggggca aatacagaca cagcaagttc 180
cttcttccct ttggaaattt ggcagctgcc ttcaccagtg agcacaaagc cacatttcaa 240
aggaaactga caaattatcc ccagctgcca gaagaagaaa tcctcactgg acggcttcct 300
gtttcctgtg gttcattatc tgattggctg cagggatgaa agtttttaag ttcataggac 360
tgatgatcct cctcacctct gcgttttcag ccggttcagg acaaagtcca atgactgtgc 420
tgtgctccat agactggttc atggtcacag tgcacccctt catgctaaac aacgatgtgt 480
gtgtacactt tcatgaacta cacttgggcc tgggttgccc cccaaaccat gttcagccac 540
acgcctacca gttcacctac cgtgttactg aatgtggcat cagggccaaa gctgtctctc 600
aggacatggt tatctacagc actgagatac actactcttc taagggcacg ccatctaagt 660
ttgtgatccc agtgtcatgt gctgcccccc aaaagtcccc atggctcacc aagccctgct 720
ccatgagagt agccagcaag agcagggcca cagcccagaa ggatgagaaa tgctacgagg 780
tgttcagctt gtcacagtcc agtcaaaggc ccaactgcga ttgtccacct tgtgtcttca 840
gtgaagaaga gcatacccag gtcccttgtc accaagcagg ggctcaggag gctcaacctc 900
tgcagccatc tcactttctt gatatttctg aggattggtc tcttcacaca gatgatatga 960
ttgggtccat gtgatcctca ggtttggggt ctcctgaaga tgctatttct agaattagta 1020
tatagtgtac aaatgtctga caaataagtg ctcttgtgac cctcatgtga gcacttttga 1080
gaaagagaaa cctatagcaa cttcatgaat taagcctttt tctatatttt tatattcatg 1140
tgtaaacaaa aaataaaata aaattctgat cgcataaaaa aaaaaaa 1187

<210> 3
<211> 1395

<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 1419725CB1

<400> 3
ctcaaaggca aacaaaagga aatgcccggc tccccatggc tgtggccagc accttcatac 60
cagggctcaa ccctcagaac cctcattata tcccagggta agactcccac ttagcgctcc 120
ccgcttgctg ctgggagtag gggtataggg tggaggaatt aagctgctta gagcatgggg 180
ccaagcggac aggttggttt gggaggctga ggtctctatg gtggtggatg gaaaaggaga 240
agcaagggac tagtgagaca ccgggtctgc cccaaagctt tgtgttgttc tctccatggg 300
aacagggtaa cagcagagga aggatcgaaa gcccaggcag agaacttaga aatgcctgct 360
ccaaatatat acactgctga gtccagaaag atcagggctt cagaggtctc cgcttctgcc 420
tctggggatg gcggggcagg gcgattctgg tgagtgggag agtctgtgct gcaggtacac 480
tggacactgc ccactacttc ggttcagcgt gggccagacc tatgggcagg tgactggtca 540
gctacttcga ggccctcctg gcctagcctg gccccctgtc caccgcacac ttctgcctcc 600
cattcggcct ccaagatctc ctgaggttcc cagggagagt ctacctgtca ggcgtgggca 660
agaaaggctc agctccagca tgatccctgg gtacacaggt ttgtacgcag gtatgcacag 720
gtgcccctcc caggagcatg tgttccaaac actaacgagt cttccttgtc cctgcctgcc 780
caggttttgt accccgggca cagttcatct ttgccaagaa ctgcagccag gtctgggccg 840
aggctctgag tgactttact cacttgcatg aaaagcaagg gagtgaagag ctaccaaagg 900
aggccaaggg aagaaaggac acagagaagg accaggtgcc agagccggag gggcagctgg 960
aggagccgac actggaggtg gtggaacaag cttctcccta ctccatggat gacagggacc 1020
ctcggaagtt cttcatgtca ggcactgcag gaatttgggc agaagcactc accaggcagt 1080
gcccaggacc ccaaacatct cccccaattt cccagaacat accctcagaa cctgggtctt 1140
ttacctaact atgggggcta cgtgccaggg tataagttcc agtttggcca cacatttggc 1200
catctcaccc atgatgctct gggcctcagc accttccaga agcagctctt ggcttaggcc 1260
actggacatc aagttccctt ccctttttcat cctatcccag ccatcctttt ggaagggaga 1320
gaggtgggtg ggagggtggg ggaacacaaa gagaaaatgg tttggaggct gagcaccttt 1380
tttattaata ggtat 1395

<210> 4
<211> 567
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2364582CB1

<400> 4
cttttctcag caaagggcag cactgctcgc tttgctgggg cagccctccc aggaagatcc 60
cagaagagat ggtctgaggt gtgtgaacac tgtgaatgtc atggacaaag ctcagcgtgg 120
gaccccatgg gacttcaggc cctgaaagca gcagatctca cctgcattgg aggatgagac 180
gccacatgaa gaagaaagaa gacgccacag acagcagcgc agtaccacct tccagagata 240
gtctcactgt gacacctagt ttggactgca atggcgcgat cttggctcac tgaaacctcc 300
gcttcccagg ttcaaacgat tctcctgcct cagccttccg agtggctggg actacaggaa 360
accagctgtc aggcctccca gatagtatca gaaagctgaa gatttccaga tcgctgcatc 420
ctaacagcaa ggtgccagaa ccctgattca actacctact gcctgttgag caactcctac 480
ataaactcct aactttattg gttacagaga cagattttag acttgtctcc catatcctgg 540
ctgcatcacc agcaataaag ccttact 567

<210> 5
<211> 920
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2737624CB1

<400> 5

gcgcaaacct gtagtcgtag ctacttggga ggctgaggca ggggaatcac ttgaaccccg 60
gaggcggagg ttgtagtggg ccgagattgt gccactgcgc tccagcgtgg gcgacagagt 120
gagactccat ctcaaaaaaa aaaaaaaaaa aatctatgct agtagattac aacttcacac 180
tagaggagtt ctggacaaag cttttaatta gtcaaactaa attaaggctc attaaaagga 240
aaggaactac tgggaaatta tgcaattcaa taatttagac tctgttacca ggatctttca 300
taaaaattta atttccataa tcataaccta aatgagttct taaagaattc tataagcaat 360
agctgattaa tgggccctgg aagatgaaga ttataactgt ttatttacct aattaaaagg 420
aaaggcagtg ccaaatatga gaggataaac aatattagtt aacatttctg ttatttatga 480
tgccaattag tagtaagata attccacagc tgtcaacttt gtttggggct ggcaacttct 540
ctgcttaaac aggctaaaag tttagtattc tgggagaagt ggctggaaga aggggtaata 600
tggtgaaagc aaattccctt tcccaggagt caagagaatt tatgtgaggt ggctcccgcc 660
tgtaatctca gcactttggg aggccaaggc aggcaatcac ctgagggcag gagaccagcc 720
tgactaacat ggagaaaccc catctctact aaaaatacaa aatgagctgg gcgtggtggt 780
gcatgcctgt aatcccagct acttgggagg ctgaggcagg agaatcgcct gaacccggga 840
ggcggaggtt gaagtgagcc aagatcgagc cactgtactc cagcctgggg gacagatcga 900
ggctgtctta aaaaaaaaaa 920

<210> 6
<211> 1564
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2867065CB1

<400> 6
aacttttaa aaattccttc tttaaatgtt tcatgattat acattgtttg caaagtaaat 60
gaaacaggtt aacttcctgg atgcttactt ttagcagaaa tatatcagat agcaaacttg 120
ctgacagatg aatcacagct ttgcccacct gtaatgagga tacactgctg gatcccagat 180
tctgtttctg gcgtctgcat acgcaagtcc ccaactcgct gaggagagct gttcagctat 240
ggctgctgtc acacattacc tgtatctttg ccagtttagc tggatgctca ttcagtctgt 300
gaatttctgg tacgtgctgg tgatgaatga tgagcacaca gagaggcgat atctgctgtt 360
tttccttctg agttggggac taccagcttt tgtggtgatt ctcctcatag ttattttgaa 420
aggaatctat catcagagca tgtcacagat ctatggactc attcatggtg acctgtgttt 480
tattccaaac gtctatgctg ctttgttcac tgcagctctt gttcctttga cgtgcctcgt 540
ggtggtgttc gtggtgttca tccatgccta ccaggtgaag ccacagtgga aagcatatga 600
tgatgtcttc agaggaagga caaatgctgc agaaattcca ctgattttat atctctttgc 660
tctgatttcc gtgacatggc tttggggagg actacacatg gcctacagac acttctggat 720
gttggttctc tttgtcattt tcaacagtct gcagggactt tatgttttca tggtttattt 780
cattttacac aaccaaatgt gttgccctat gaaggccagt tacactgtgg aaatgaatgg 840
gcatcctgga cccagcacag ccttttttcac gcccgggagt ggaatgcctc ctgctggagg 900
ggaaatcagc aagtccaccc agaatctcat cggtgctatg gaggaggtgc cacctgactg 960
ggagagagca tccttccaac agggcagtca ggccagccct gatttaaagc caagtccaca 1020
aaatggagcc acgttcccgt cctctggagg atatggccag gggtcactga tagccgatga 1080
ggagtcccag gagtttgatg atttaatatt tgcattaaaa actggtgctg gtctcagtgt 1140
cagtgataat gaatctggtc aaggcagcca ggaggggggc accttgactg actcccagat 1200
cgtggagctc aggaggatac ccatcgccga cactcacctg tagcacctca ctaaccattc 1260
gactgagcac actttcatat ttgtatcagc ttttgtgcta aaactctcta agtacatcca 1320
cctgtgtaat aggaacctgt gaattgtact ggatgattaa tacaaacgtg attgttgtat 1380
ttggagtata aattactgat tgtatgtgac ctgaaaattc actgctataa gaaaggtgga 1440
gtcagtttgt atcagttaat aggatgttca tattccaagg atattagttg ttttttaat 1500
catcctatat ggctaacatt gtttaatgaa agtaataatc aataaagcaa tagaatctaa 1560
aaaa 1564

<210> 7
<211> 752
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2961563CB1

<400> 7
gatgggaggg ggtagggagg atgaaaaaat caccattacc acctgaagct tttccattct 60
tgttttaata gcattgatga tttgcaccat ttggagatag agctctggct tatttattat 120
aaaacaggag ctgggcacag tggtgcctac ccacaatgcc agctactctg aaggctgagg 180
tgggagaatt gctggagggc aattcgaaac cagtctgggc aagatagtga gacctcatct 240
caaaaaaaaa aaaagaagga aaaaacaaag gagttgatga aaccaaatta tgacaaacta 300
aattaaaaaa ttaaaaaata gcattagtgg gaacactgtg gtatttcctg aggtattaaa 360
attttacaac tagtaaaggt atttattata atttattcgg attgtttctt agtttatttt 420
tgacatggat aaaattattg tttatctagg agcctcagag ttgagagagg atagttgtag 480
agaagtattt ctaaaatgga attatcatcc actaccagct aagagaaaat agttacttgt 540
ttaaagatac cttaaagcaa catatattgt agtttcaaga gaacaggctg acagtcaaga 600
aataggcttt taacaaaccc actggtttgc tggattctaa ttgcctcctt caaactatgt 660
aattgtgtct ttctggtttc tgtggagata ttttcccact tgtttaattt tacactggaa 720
taaaacaagt tatcactgtt aaaaaaaaaa aa 752

<210> 8
<211> 212
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 065781CD1

<400> 8
Met Lys Val Phe Lys Phe Ile Gly Leu Met Ile Leu Leu Thr Ser
1 5 10 15
Ala Phe Ser Ala Gly Ser Gly Gln Ser Pro Met Thr Val Leu Cys
20 25 30
Ser Ile Asp Trp Phe Met Val Thr Val His Pro Phe Met Leu Asn
35 40 45
Asn Asp Val Cys Val His Phe His Glu Leu His Leu Gly Leu Gly
50 55 60
Cys Pro Pro Asn His Val Gln Pro His Ala Tyr Gln Phe Thr Tyr
65 70 75
Arg Val Thr Glu Cys Gly Ile Arg Ala Lys Ala Val Ser Gln Asp
80 85 90
Met Val Ile Tyr Ser Thr Glu Ile His Tyr Ser Ser Lys Gly Thr
95 100 105
Pro Ser Lys Phe Val Ile Pro Val Ser Cys Ala Ala Pro Gln Lys
110 115 120
Ser Pro Trp Leu Thr Lys Pro Cys Ser Met Arg Val Ala Ser Lys
125 130 135
Ser Arg Ala Thr Ala Gln Lys Asp Glu Lys Cys Tyr Glu Val Phe
140 145 150
Ser Leu Ser Gln Ser Ser Gln Arg Pro Asn Cys Asp Cys Pro Pro
155 160 165
Cys Val Phe Ser Glu Glu Glu His Thr Gln Val Pro Cys His Gln
170 175 180
Ala Gly Ala Gln Glu Ala Gln Pro Leu Gln Pro Ser His Phe Leu
185 190 195
Asp Ile Ser Glu Asp Trp Ser Leu His Thr Asp Asp Met Ile Gly
200 205 210
Ser Met

<210> 9
<211> 334
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No.: 2867065CD1

<400> 9

```
Met Ala Ala Val Thr His Tyr Leu Tyr Leu Cys Gln Phe Ser Trp
 1               5                  10                  15
Met Leu Ile Gln Ser Val Asn Phe Trp Tyr Val Leu Val Met Asn
            20                  25                  30
Asp Glu His Thr Glu Arg Arg Tyr Leu Leu Phe Phe Leu Leu Ser
        35                  40                  45
Trp Gly Leu Pro Ala Phe Val Val Ile Leu Leu Ile Val Ile Leu
    50                  55                  60
Lys Gly Ile Tyr His Gln Ser Met Ser Gln Ile Tyr Gly Leu Ile
65                  70                  75
His Gly Asp Leu Cys Phe Ile Pro Asn Val Tyr Ala Ala Leu Phe
                80                  85                  90
Thr Ala Ala Leu Val Pro Leu Thr Cys Leu Val Val Val Phe Val
            95                  100                 105
Val Phe Ile His Ala Tyr Gln Val Lys Pro Gln Trp Lys Ala Tyr
        110                 115                 120
Asp Asp Val Phe Arg Gly Arg Thr Asn Ala Ala Glu Ile Pro Leu
    125                 130                 135
Ile Leu Tyr Leu Phe Ala Leu Ile Ser Val Thr Trp Leu Trp Gly
140                 145                 150
Gly Leu His Met Ala Tyr Arg His Phe Trp Met Leu Val Leu Phe
155                 160                 165
Val Ile Phe Asn Ser Leu Gln Gly Leu Tyr Val Phe Met Val Tyr
            170                 175                 180
Phe Ile Leu His Asn Gln Met Cys Cys Pro Met Lys Ala Ser Tyr
        185                 190                 195
Thr Val Glu Met Asn Gly His Pro Gly Pro Ser Thr Ala Phe Phe
    200                 205                 210
Thr Pro Gly Ser Gly Met Pro Pro Ala Gly Gly Glu Ile Ser Lys
215                 220                 225
Ser Thr Gln Asn Leu Ile Gly Ala Met Glu Glu Val Pro Pro Asp
            230                 235                 240
Trp Glu Arg Ala Ser Phe Gln Gln Gly Ser Gln Ala Ser Pro Asp
        245                 250                 255
Leu Lys Pro Ser Pro Gln Asn Gly Ala Thr Phe Pro Ser Ser Gly
    260                 265                 270
Gly Tyr Gly Gln Gly Ser Leu Ile Ala Asp Glu Glu Ser Gln Glu
275                 280                 285
Phe Asp Asp Leu Ile Phe Ala Leu Lys Thr Gly Ala Gly Leu Ser
            290                 295                 300
Val Ser Asp Asn Glu Ser Gly Gln Gly Ser Gln Glu Gly Gly Thr
        305                 310                 315
Leu Thr Asp Ser Gln Ile Val Glu Leu Arg Arg Ile Pro Ile Ala
    320                 325                 330
Asp Thr His Leu
```